Accelerated Publication

Protein *trans*-Splicing and Cyclization by a Naturally Split Intein from the *dnaE* Gene of *Synechocystis* Species PCC6803*

Thomas C. Evans, Jr., Deana Martin, Reto Kolly‡, Daniel Panne§, Luo Sun, Inca Ghosh, Lixin Chen, Jack Benner, Xiang-Qin Liu¶, and Ming-Qun Xu‖

From New England Biolabs, Inc., Beverly, Massachusetts 01915, §Biozentrum, University of Basel, CH-4056 Basel, Switzerland, and ¶Department of Biochemistry and Molecular Biology, Dalhousie University, Halifax, Nova Scotia B3H 4H7, Canada

A naturally occurring split intein from the *dnaE* gene of *Synechocystis* sp. PCC6803 (*Ssp* DnaE intein) has been shown to mediate efficient *in vivo* and *in vitro* trans-splicing in a foreign protein context. A *cis*-splicing *Ssp* DnaE intein construct displayed splicing activity similar to the *trans*-splicing form, which suggests that the N- and C-terminal intein fragments have a high affinity interaction. An *in vitro* trans-splicing system was developed that used a bacterially expressed N-terminal fragment of the *Ssp* DnaE intein and either a bacterially expressed or chemically synthesized intein C-terminal fragment. Unlike artificially split inteins, the *Ssp* DnaE intein fragments could be reconstituted *in vitro* under native conditions to mediate splicing as well as peptide bond cleavage. This property allowed the development of an on-column *trans*-splicing system that permitted the facile separation of reactants and products. Furthermore, the *trans*-splicing activity of the *Ssp* DnaE intein was successfully applied to the cyclization of proteins *in vivo*. Also, the isolation of the unspliced precursor on chitin resin allowed the cyclization reaction to proceed *in vitro*. The *Ssp* DnaE intein thus represents a potentially important protein for *in vivo* and *in vitro* protein manipulation.

Protein splicing elements, termed inteins (1), catalyze their own excision from a primary translation product with the concomitant ligation of the flanking protein sequences (reviewed in Refs. 2–4). Inteins catalyze three highly coordinated reactions at the N- and C-terminal splice junctions (5, 6): 1) an acyl rearrangement at the N-terminal cysteine or serine; 2) a trans-esterification reaction between the two termini to form a branched ester or thioester intermediate; and 3) peptide bond cleavage coupled to cyclization of the intein C-terminal asparagine to free the intein. Inteins have been engineered to be versatile tools in protein purification (7–13), protein ligation (9,

10, 12, 14–18), and in the formation of cyclic proteins and peptides (11, 19, 20). However, the ligation and cyclization approaches were limited by the need to generate an N-terminal cysteine and/or C-terminal thioester intermediate *in vitro*.

In addition to inteins engineered to *trans*-splice (21–24), a naturally occurring split intein was recently identified in the *dnaE* gene encoding the catalytic subunit of DNA polymerase III of *Synechocystis* sp. PCC6803 (25). The N-terminal half of DnaE, followed by a 123-amino acid intein sequence, and the C-terminal half, preceded by a 36-amino acid intein sequence, are encoded by two open reading frames located more than 745 kilobases apart in the genome. When co-expressed in *Escherichia coli*, the two DnaE-intein fragments exhibited protein *trans*-splicing (25). In this report we have further investigated the cis- and *trans*-splicing activities of the *Ssp* DnaE intein in a foreign protein context. Furthermore, novel methods were developed that allow the on-column ligation of protein fragments as well as the *in vivo* and *in vitro* cyclization of proteins by intramolecular *trans*-splicing of the *Ssp* DnaE intein fragments.

EXPERIMENTAL PROCEDURES

Construction of Plasmids—pMEB1 was constructed by replacing the *Sce* VMA intein in pMYB129 (7) with the *Ssp* DnaE intein sequence spanning residues 5–123 to create a fusion gene composed of *E. coli* maltose-binding protein (MBP)[1] (26), the *Ssp* DnaE intein (residues 5–123), and the *Bacillus circulans* chitin binding domain (CBD) (27). The *Ssp* DnaE intein fragment was amplified from plasmid pDnaE-C-209 with primers 5'-TTTGGTACCGAAATTTTAACCGTTGAG-3' and 5'-GGCTCTTCCTTTAATTGTCCCAGCGTCAAG-3'. The N-terminal splice junction sequence, containing the flanking 5 native N-extein residues and the 5 intein N-terminal residues, was inserted between MBP and the intein coding regions by linker insertion into the *Xho*I and *Kpn*I sites in pMEB1 to create pMEB2. The linker was formed by annealing oligonucleotides, 5'-TCGAGAAATTTGCTGAATATTGCCTGTCTTTTG-GTAC-3' and 5'-CAAAAGACAGGCAATATTCAGCAAATTTC-3'.

The DNA sequence encoding the C-terminal 36 amino acid residues and the first 3 C-extein residues (5'-ATGGTTAAAGTTATCGGTCGT-AGATCTCTGGGCGTGCAGCGCATCTTTGATATCGGTCTGCCCGCA-GGACCATAACTTTCTGCTAGCCAACGGCGCTATCGCTGCTAACTG-CTTTAACAAATCC-3') was inserted into pMEB2 to create pMEB3, which expresses a fusion protein (MEB) composed of MBP, the full-length *Ssp* DnaE intein (residues 1–159) with 5 native extein residues at its N terminus and 3 native residues at its C terminus, and the CBD. A translation termination codon was introduced into pMEB2 following the codon for Lys[123] of the *Ssp* DnaE intein by insertion of a linker formed by annealing oligonucleotides 5'-AAATAAGGAGGTTAATAAA-AGGAAGAGCCATGGCGCGCCCTTAATTAAA-3' and 5'-CCGGTTTAA-TTAAGGCGCGCCATGGCTCTTCCTTTTATTAACCTCCTTA-3'. The resulting plasmid, pMEB4, expresses a fusion protein composed of MBP and the N-terminal 123 residues of the *Ssp* DnaE intein (DnaE(N)). pKEB1 contains the kanamycin resistance gene and the p15a origin of replication from pACYC177 (28). It also expresses a fusion protein composed of the 36 C-terminal amino acids of the *Ssp* DnaE intein (DnaE(C)) followed by 3 native extein residues and the CBD. pBEL11 expresses a CBD-DnaE(C)-T4 DNA ligase fusion protein in the pBSL-C155 vector (10).

pMEB8 was generated by transferring the 0.6-kilobase *Xho*I to *Pst*I fragment of pMEB3 into pMYB5 (NEB). Mutation of the extein residues in pMEB8 was performed by linker substitutions using the *Xho*I and *Kpn*I sites flanking the N-terminal splice junction or the *Nhe*I and *Age*I

‡ Present address: Dept. of Biochemistry, University of Zurich, CH-8057 Zurich, Switzerland.

‖ To whom correspondence should be addressed: New England Biolabs, Inc., 32 Tozer Rd., Beverly, MA 01915-5599. Tel.: 978-927-5054; Fax: 978-921-1350; E-mail: xum@neb.com.

[1] The abbreviations used are: MBP, maltose-binding protein; *Ssp* DnaE intein, a naturally split intein from the *dnaE* gene of *Synechocystis* sp. PCC6803; DnaE(N), the N-terminal 123 amino acid residues of the *Ssp* DnaE intein; DnaE(C), the C-terminal 36 amino acid residues of the *Ssp* DnaE intein; CBD, chitin binding domain; FXa, factor Xa; PAGE, polyacrylamide gel electrophoresis.
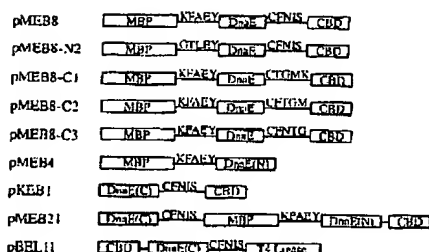
Fig. 1. *Ssp* DnaE intein *cis-* and *trans-*splicing constructs. The *cis-*splicing constructs, pMEB8, pMEB8-N2, pMEB8-C1, -C2, and -C3, all use MBP and the CBD as the N- and C-exteins, respectively. The differences are in the extein residues adjacent to the intein and are represented by their single letter code for ease of comparison. The constructs used in the two-plasmid, *trans-*splicing system were pMEB4 and pKEB1, which contain the N- and C-terminal *Ssp* DnaE intein fragments, DnaE(N) and DnaE(C), respectively. The intramolecular *trans-*splicing construct, pMEB21, placed DnaE(C) and DnaE(N) at the N and C terminus of MBP, respectively.
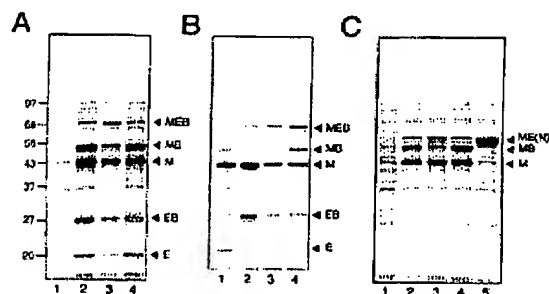

FIG. 2. **Examination of *in vivo* splicing of the *Ssp* DnaE intein by SDS-PAGE.** A, *cis-*splicing activity of the *Ssp* DnaE intein investigated by expression of the MBP-DnaE-CBD fusion (*MEB*) protein in *E. coli*. *Lane 1*, uninduced crude cell extract; *lane 2*, crude cell extract following induction at 15 °C overnight. The MEB precursor (68 kDa), the splicing products MB (50 kDa) and E (18 kDa), and the cleavage products M (43 kDa) and EB (25 kDa) are visible. *Lane 3*, crude cell extract after a 2-h induction at 37 °C; *lane 4*, crude cell extract after a 2-h induction at 37 °C followed by overnight incubation at 15 °C. B, the *cis-*splicing activity of the *Ssp* DnaE intein *in vivo* with mutated extein residues. *Lane 1*, *in vivo* splicing activity with 2 native N-extein and 3 native C-extein residues (pMEB-N2). The splicing of the *Ssp* DnaE intein with 5 native N-extein residues and 1 (pMEB8-C1, *lane 2*), 2 (pMEB8-C2, *lane 3*), or 3 native C-extein residues (pMEB8-C3, *lane 4*) is also shown. C, the *Ssp* DnaE intein *in vivo* *trans-*splicing activity investigated by co-expression of the MBP-DnaE(N) (57 kDa) and the DnaE(C)-CBD (10 kDa) fusion proteins. *Lane 1*, uninduced crude cell extract. Crude cell extract after induction of protein expression at 15 °C (*lane 2*), at 37 °C (*lane 3*), and at 37 °C followed by incubation with shaking at 15 °C (*lane 4*) all displayed precursor (ME(N)), spliced product (MB) and cleavage product (M). *Lane 5*, induction of protein expression with ME(N), but not E(C)B, displayed no detectable splicing or cleavage. All samples were analyzed by Coomassie Blue-stained 12% SDS-PAGE gels.

sites flanking the C-terminal splice junction. pMEB8-N2 retains 2 native N-extein residues whereas pMEB8-C1, -C2, or -C3 possess 1, 2, or 3 native C-extein residues (Fig. 1). The intramolecular *trans-*splicing construct, pMEB21, expresses a fusion protein with the DnaE(C) immediately followed by amino acid residues CFNISTG, MBP, which terminates with amino acid residues GTLEKFAEY and then DnaE(N)-CBD.

Ssp DnaE cis- and trans-Splicing in Vivo—E. coli ER2566 cells (7) bearing pMEB8 or two compatible plasmids, pMEB4 and pKEB1, were grown in LB medium containing the appropriate antibiotic selection at 37 °C to an A_{600} of 0.5. Protein expression was induced by the addition of 0.3 mM isopropyl β-D-thiogalactopyranoside at either 15 °C for 16 h or at 37 °C for 2 h. Crude cell extracts were visualized by electrophoresis on a 12% Tris/glycine gel (Novex, San Diego, CA) followed by staining with Coomassie Brilliant Blue.

Protein Purification—ER2566 cells containing pMEB2 or pBEL11 were grown at 37 °C to an A_{600} of 0.5. Following isopropyl β-D-thiogalactopyranoside (0.5 mM) induced protein expression overnight at 15 °C, cells were harvested by centrifugation at 3,000 × g for 30 min. The
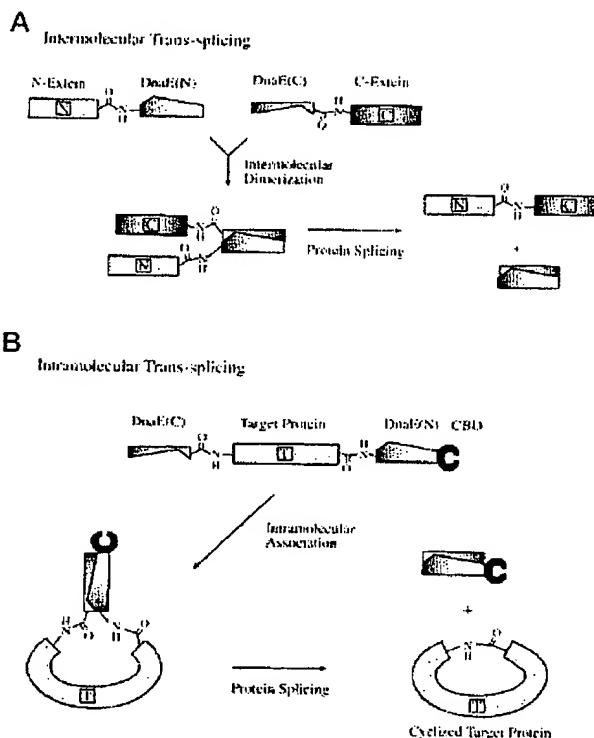

FIG. 3. **Protein *trans-*splicing and cyclization reactions using the *Ssp* DnaE intein.** A, intermolecular *trans-*splicing. The association of the N-terminal and C-terminal *Ssp* DnaE intein fragments, DnaE(N) and DnaE(C), respectively, aligns the two splice junctions for the fusion of the N- and C-extein sequences. The splicing reaction presumably occurs via the same splicing pathway as the *cis-*splicing pathway proposed previously (5, 6). Cleavage at the N-terminal splice junction can occur by hydrolysis or nucleophilic attack of the thioester bond formed at the C terminus of the N-extein. B, intramolecular *trans-*splicing. A target protein is sandwiched between the intein C-terminal segment (36 amino acids) and the intein N-terminal segment (123 amino acids). Splicing joins the N terminus of the target protein to its own C terminus through a peptide bond. The presence of a CBD fused to the C terminus of the intein N-terminal segment facilitated purification of the precursor protein and the subsequent *in vitro* cyclization reaction on chitin resin.

MBP-DnaE(N)-CBD (ME(N)B) fusion protein was purified by amylose as described previously (9). The cell pellet was resuspended in Buffer A (20 mM Tris-HCl, pH 7.0, containing 500 mM NaCl) and lysed by sonication. After centrifugation at 23,000 × g for 30 min the supernatant was applied to a 15-ml amylose resin (NEB) equilibrated in Buffer A. The resin was washed with 10–15 column volumes of Buffer A. The fusion protein was eluted with Buffer B (20 mM Tris-HCl, pH 7.0, containing 500 mM NaCl and 10 mM maltose). Protein concentrations were determined using the Bio-Rad protein assay (Bio-Rad).

*In Vitro trans-Splicing and/or Cleavage Assay—trans-*Splicing and/or *trans-*cleavage studies of the *Ssp* DnaE intein were conducted *in vitro* using the purified ME(N)B and two 40-amino acid peptides, synthesized as described previously (9), consisting of the C-terminal 36 amino acids of the *Ssp* DnaE intein, with either an Asn (Splice-pep) or an Ala at residue 36 (Cleav-pep), and the next four naturally occurring amino acids (CFNK). The splicing peptide had a biotinylated lysine (K*) as the C-terminal residue. The reaction consisted of adding either the splicing or cleavage peptide (500 μM final concentration) to ME(N)B (1 mg/ml) in reaction buffer (100 mM Tris-HCl, pH 7.0, containing 500 mM NaCl) followed by incubation overnight at room temperature. The on-column *trans-*splicing used the CBD-DnaE(C)-T4 DNA ligase protein absorbed onto a chitin resin in which unbound protein was washed off with 20 column volumes of Buffer A. The ME(N)B fusion protein (9 μM), either free in solution or prebound to chitin beads, was then added to the chitin-bound CBD-DnaE(C)-T4 DNA ligase (3 μM). The reactions were then incubated for 16 h at the appropriate temperature in Buffer A and monitored by SDS-PAGE.
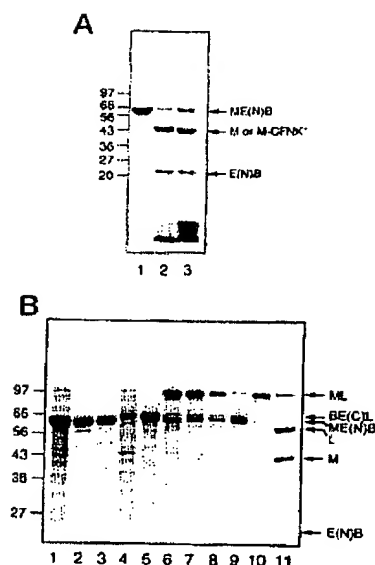
FIG. 4. In vitro trans-splicing of the *Ssp* DnaE intein. *A*, 10–20% SDS-PAGE gel of peptide-induced splicing and cleavage of MBP-DnaE(N)-CBD (*ME(N)B*). *Lane 1*, amylose-purified ME(N)B (64 kDa). *Lane 2*, ME(N)B, 1 mg/ml, following overnight incubation with the splicing peptide (Splice-pep, 500 μM). The cleavage product (M) and spliced product (M-CFNK*) co-migrate at 43 kDa. *Lane 3*, ME(N)B, 1 mg/ml, after overnight incubation with the cleavage peptide (Cleav-pep, 500 μM). Both Splice-pep and Cleav-pep are based on the C-terminal fragment of the *Ssp* DnaE intein as described under "Experimental Procedures." *B*, *trans*-splicing of MBP-DnaE(N)-CBD (*ME(N)B*) and CBD-DnaE(C)-T4 ligase (*BE(C)L*) examined by 12% SDS-PAGE. *Lane 1*, crude cell extract after induction of ME(N)B expression. *Lane 2*, ME(N)B following purification over an amylose column; *lane 3*, ME(N)B that was bound to chitin beads and eluted with SDS. The CBD permits binding to a chitin resin. *Lane 4*, crude cell extract after induction of BE(C)L expression; *lane 5*, chitin-bound BE(C)L that was removed by treatment with SDS; *lane 6*, incubation of amylose-purified ME(N)B and chitin-bound BE(C)L at 4 °C for 16 h followed by elution of the chitin resin with SDS. The fusion proteins ME(N)B and BE(C)L were bound to separate batches of chitin beads, and the chitin-bound proteins were mixed followed by elution of the beads with SDS after incubation at 4 °C (*lane 7*), 16 °C (*lane 8*), or 37 °C (*lane 9*) for 16 h. *Lane 10*, the supernatant from the chitin bead mixture described in *lane 8*. Note that the spliced product (ML) is free in solution whereas the reactants remain bound to the chitin beads. *Lane 11*, Fxa treatment (1:100, FXa:ML) of the supernatant fraction. *M*, MBP (43 kDa); *L*, T4 DNA ligase (58 kDa).

Protein Cyclization and Analysis—ER2566 cells bearing pMEB21 were grown, induced, harvested, and lysed as described under "Protein Purification." The clarified supernatant from cells induced at 15 °C was applied to an amylose resin (10- ml bed volume) whereas the clarified supernatant from cells induced at 37 °C was applied to a chitin resin (15-ml bed volume). Unbound proteins were washed from the resin with 20 column volumes of Buffer A. Proteins were eluted from the amylose column with Buffer B. The intramolecular *trans*-splicing reaction proceeded *in vitro* when the chitin column was incubated for 20 h at room temperature. Reaction products were eluted from the resin with Buffer A. The cyclic MBP was analyzed by treatment with FXa (1:100, FXa:protein mass to mass ratio) overnight at 4 °C to generate linearized MBP. The proteolyzed proteins were subjected to amino acid sequencing using a Procise 494 protein sequencer (PE Applied Biosystems, Foster City, CA).

RESULTS

In Vivo Splicing Activity of the Ssp DnaE Intein—The *cis*-splicing activity of the *Ssp* DnaE intein was investigated by expression of MEB in *E. coli* cells bearing pMEB8 (Fig. 1). The presence of the MBP-CBD (MB) band following induction of protein expression demonstrated that the *Ssp* DnaE intein can splice in *cis* with only 5 native N-terminal and 3 native C-terminal extein residues (Fig. 2A, *lane 2*). The identity of splicing products was confirmed by Western blot analysis using
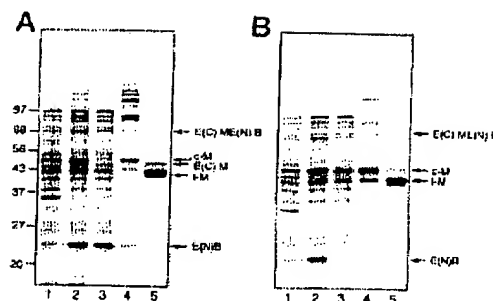

FIG. 5. Protein circularization in *E. coli* and on a chitin column. *A*, *in vivo* protein cyclization. *Lane 1*, uninduced crude cell extract; *lane 2*, crude cell extract following induction at 15 °C contains the precursor DnaE(C)-MBP-DnaE(N)-CBD (*E(C)ME(N)B*, 65 kDa), cyclic MBP (*c-M*, 47 kDa), linear MBP (*l-M*, 43 kDa), the DnaE(C)-MBP (*E(C)M*, 45 kDa), and DnaE(N)-CBD (*E(N)B*, 23 kDa). *Lane 3*, clarified cell extract from *lane 2* following passage over an amylose column. Note that the cyclic MBP binds to amylose. *Lane 4*, proteins eluted from amylose resin; *lane 5*, the eluted sample incubated with factor Xa (1:100, FXa:MBP). FXa treatment also resulted in the release of a 45-kDa species corresponding to E(C)M. *B*, *in vitro* protein circularization. *Lane 1*, uninduced cell extract; *lane 2*, crude cell extract following induction at 37 °C; *lane 3*, clarified cell extract following passage over a chitin column; *lane 4*, proteins eluted from the chitin column following incubation at 23 °C for 16 h; *lane 5*, incubation of the chitin-eluted sample with factor Xa (1:100, FXa:MBP). All reactions were performed as described under "Experimental Procedures" and were analyzed on a 12% SDS-PAGE gel.

anti-MBP and anti-CBD antibodies and binding to chitin and amylose resins (data not shown). In addition to the spliced product, there was significant cleavage of the peptide bond at the N terminus of the *Ssp* DnaE intein.

The role of the extein amino acid residues was investigated by mutation of the distal extein residues in the *cis*-splicing construct (Fig. 2B). Splicing products were detected in mutants with either 2 proximal N-extein residues or 3 proximal C-extein residues (Figs. 1 and 2B, *lanes 1* and *4*). However, reduction of the C-extein sequence to 1 or 2 native amino acid residues inhibited splicing (Fig. 2B, *lanes 2* and *3*).

Interestingly, the same pattern of splicing and cleavage as seen with the *cis*-splicing construct was observed for in vivo *trans*-splicing (Fig. 2C, *lane 2*). Furthermore, both the pMEB4 and pKEB1 plasmids were necessary to induce processing of the precursor protein (Fig. 2C, *lane 5*). There was a slight accumulation of *Ssp* DnaE intein precursor protein when protein expression was induced at 37 °C, and this was processed after further growth overnight at 15 °C (Fig. 2, *A* and *C*).

In Vitro trans-Splicing with the Ssp DnaE Intein—The *in vitro trans*-splicing (Fig. 3) and/or *trans*-cleavage activity of the *Ssp* DnaE intein was studied using the bacterially expressed ME(N)B precursor and two peptides, Splice-pep and Cleav-pep, that mimic the C-terminal *Ssp* DnaE intein fragment (see "Experimental Procedures"). Both the Splice-pep and the Cleav-pep could activate ME(N)B, resulting in bands corresponding to the expected spliced and/or cleavage product (Fig. 4A). Furthermore, the ME(N)B precursor was stable in the absence of either peptide (Fig. 4A, *lane 1*). The cleavage and splicing products, MBP and MBP-CFNK*, respectively, are indistinguishable by SDS-PAGE. However, a Western blot using anti-biotin antibody indicated that splicing was occurring, albeit the extent of reaction could not be determined (data not shown). Efficient *in vitro trans*-splicing occurred between two bacterially expressed proteins, MBP-DnaE(N)-CBD and CBD-DnaE(C)-T4 DNA ligase, yielding spliced product, MBP-T4 DNA ligase (ML), at 4 and 16 °C but significantly less at 37 °C (Fig. 4B). Interestingly, little difference in splicing efficiency was observed when either chitin-bound or free ME(N)B was

used to react with the chitin-bound BE(C)L. Perhaps the chitin binding domain, although possessing high affinity to chitin, can dissociate and reassociate with the resin and allow the two intein fragments to react. Following FXa proteolysis of the released ligation product, amino acid sequencing of the 58-kDa band (expected for T4 DNA ligase) yielded NH$_2$-GTLEKFAEY-CFNIST-COOH, which corresponds to the expected sequence of the ligation junction.

In Vivo/in Vitro Protein Cyclization with the Ssp DnaE In-tein—To test the feasibility of protein cyclization by the *Ssp* DnaE intein, the DnaE(C) and DnaE(N) fragments were fused to the N and C termini of MBP, respectively (pMEB21, Figs. 1, 3, and 5). SDS-PAGE (Fig. 5A, *lane 2*) and Western blot analysis (data not shown) of crude cell extract of cells expressing pMEB21 showed that there was precursor protein, DnaE(C)-MBP-DnaE(N)-CBD, linear MBP, circular MBP, DnaE(N)-CBD, and DnaE(C)-MBP. The putative linear and cyclic MBP species as well as higher molecular weight species (Fig. 5A, *lane 4*) were found to bind to amylose and elute with maltose. The maltose eluted proteins were subjected to FXa proteolysis and amino acid sequencing. The upper portion of the 43-kDa band yielded NH$_2$-GTLEKFAEYXFNISTGM-COOH, which matched the sequence for the cyclic MBP that was linearized with FXa. Sequencing the lower part of the 43-kDa band gave NH$_2$-XFNISTGM-COOH, which matched the N terminus of the linear MBP, which had not undergone cyclization. NH$_2$-XVKIGRRSLGV-COOH was obtained from the 45-kDa band and correlates with the expected sequence from the DnaE(C)-MBP product. The *X* designates a sequencing cycle in which no amino acid could be assigned with confidence.

Precursor consisting of DnaE(C)-MBP-DnaE(N)-CBD could be obtained by inducing protein expression for 2 h at 37 °C (Fig. 5B, *lane 2*). The precursor was immobilized on a chitin resin through the CBD, and intramolecular *trans*-splicing proceeded overnight at 23 °C. Fractions from the chitin resin contained cyclic and linear MBP species (Fig. 5B, *lane 4*). FXa treatment of the isolated proteins (Fig. 5B, *lane 5*) followed by amino acid sequencing confirmed the presence of both the linear and circular forms.

DISCUSSION

The present study demonstrated that the *Ssp* DnaE intein was capable of splicing in a non-native protein context and determined that the intein fragments can self-associate with no more than 5 native extein residues. This implies that the interaction of the two intein halves in the natural condition is at least partly, or perhaps entirely, dominated by the association of the intein fragments and not by the extein residues. However, the presence of more extein residues may increase the efficiency of the splicing reaction and be vital to the effectiveness of this intein in the DnaE protein. Interestingly, the *trans*- and *cis*-splicing activities of the *Ssp* DnaE intein were almost identical; this indicates that the N- and C-terminal domains of this intein have a high affinity interaction.

Inteins represent a unique opportunity to perform protein manipulations *in vivo* as well as *in vitro*. The present work has demonstrated the facile production of circular and possibly multimeric proteins in *E. coli* and opens up new avenues to produce stable, bioactive proteins and peptides in living cells. Also, a recent paper, published following completion of this manuscript, describes the *in vivo* cyclization of proteins using the *Ssp* DnaE intein, which they term the *in vivo* split intein-mediated circular ligation of peptide and proteins (SICLOPPS) (29). We propose that the term SICLOPPS be used in the future to describe the *in vivo* cyclization reaction. *trans*-Splicing allows the study of cyclic peptides and proteins in a cellular environment. Furthermore, the possibility of assembling multimeric proteins *in vivo* is a novel and potentially useful technology.
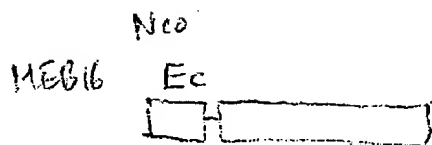
The *Ssp* DnaE intein may be the intein of choice for *in vitro* *trans*-splicing experiments as it has been demonstrated to undergo the *trans*-splicing reaction without the need for a denaturation/renaturation step, as was necessary with other inteins (21, 23, 24). Also, the addition of the CBD to either the N- or C-terminal intein fragment had no detectable inhibitory effect on splicing. These unique properties allowed fusion proteins of both the intein N- and C-terminal fragments to be immobilized on a chitin resin for the subsequent ligation of two proteins under native conditions (Fig. 4B). This represents a significant advantage over other intein-based protein fusion techniques because the spliced products are easily purified away from the column-bound reactants. Furthermore, association of the two halves of the *Ssp* DnaE intein permits *trans*-splicing to occur at relatively low concentrations of reactants.

The present work demonstrated the use of inteins in both the *in vivo* and the *in vitro* manipulation of proteins. In particular, the *Ssp* DnaE intein represents an interesting and useful protein that *trans*-splices effectively *in vivo* with only minimal extein residue sequence and does not require cumbersome denaturation/renaturation steps for use *in vitro*. Future work should refine procedures to build and modify proteins in a cellular environment in much the same way as it is now possible to perform *in vitro*.
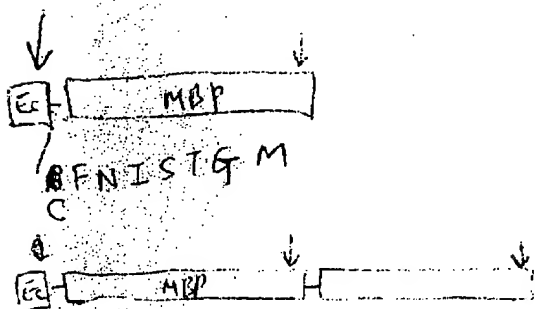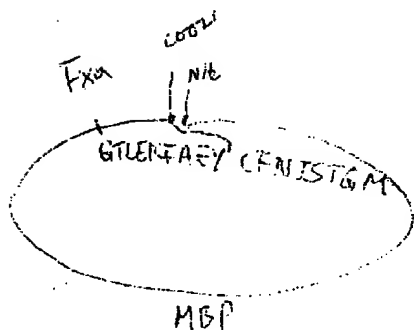
REFERENCES

1. Perler, F. B., Davis, E. O., Dean, G. E., Gimble, F. S., Jack, W. E., Neff, N., Noren, C. J., Thorner, J., and Belfort, M. (1994) *Nucleic Acids Res.* 22, 1125–1127
2. Paulus, H. (1998) *Chem. Soc. Rev.* 27, 375–386
3. Perler, F. B. (1998) *Cell* 92, 1–4
4. Shao, Y., and Kent, S. B. (1997) *Chem. Biol.* 4, 187–194
5. Xu, M.-Q., and Perler, F. B. (1996) *EMBO J.* 15, 5146–5153
6. Chong, S., Shao, Y., Paulus, H., Bonner, J., Perler, F. B., and Xu, M.-Q. (1996) *J. Biol. Chem.* 271, 22159–22168
7. Chong, S., Mersha, F. B., Comb, D. G., Scott, M. E., Landry, D., Vence, L. M., Perler, F. B., Benner, J., Kucera, R. B., Hirvonen, C. A., Pelletier, J. J., Paulus, H., and Xu, M.-Q. (1997) *Gene (Amst.)* 192, 271–281
8. Chong, S., Montello, G. E., Zhang, A., Cantor, E. J., Liao, W., Xu, M.-Q., and Benner, J. (1998) *Nucleic Acids Res.* 26, 5109–5115
9. Evans, T. C., Jr., Benner, J., and Xu, M.-Q. (1998) *Protein Sci.* 7, 2256–2264
10. Mathys, S., Evans, T. C., Jr., Chute, I. C., Wu, H., Chong, S., Benner, J., Liu, X.-Q., and Xu, M.-Q. (1999) *Gene (Amst.)* 231, 1–13
11. Evans, T. C., Jr., Benner, J., and Xu, M.-Q. (1999) *J. Biol. Chem.* 274, 18359–18363
12. Southworth, M. W., Amaya, K., Evans, T. C., Jr., Xu, M.-Q., and Perler, F. B. (1999) *BioTechniques* 27, 110–120
13. Wood, D. W., Wu, W., Belfort, G., Derbyshire, V., and Belfort, M. (1999) *Nature Biol Technology* 17, 889–892
14. Evans, T. C., Jr., Benner, J., and Xu, M.-Q. (1999) *J. Biol. Chem.* 274, 3923–3926
15. Cotton, G. J., Ayers, B., Xu, R., and Muir, T. W. (1999) *J. Am. Chem. Soc.* 121, 1100–1101
16. Muir, T. W., Sondhi, D., and Cole, P. A. (1998) *Proc. Natl. Acad. Sci. U. S. A.* 95, 6705–6710
17. Severinov, K., and Muir, T. W. (1998) *J. Biol. Chem.* 273, 16205–16209
18. Xu, R., Ayers, B., Cowburn, D., and Muir, T. W. (1999) *Proc. Natl. Acad. Sci. U. S. A.* 96, 388–393
19. Iwai, H., and Pluckthun, A. (1999) *FEBS Lett.* 459, 166–172
20. Camarero, J. A., and Muir, T. W. (1999) *J. Am. Chem. Soc.* 121, 5597–5598
21. Yamazaki, T., Otomo, T., Oda, N., Kyoguku, Y., Uegaki, K., Ito, N., Ishino, Y., and Nakamura, H. (1998) *J. Am. Chem. Soc.* 120, 5591–5592
22. Wu, H., Xu, M.-Q., and Liu, X.-Q. (1998) *Biochim. Biophys. Acta* 1387, 422–432
23. Mills, K. V., Lew, B. M., Jiang, S., and Paulus, H. (1998) *Proc. Natl. Acad. Sci. U. S. A.* 95, 3543–3548
24. Southworth, M. W., Adam, E., Panne, D., Byer, R., Kautz, R., and Perler, F. B. (1998) *EMBO J.* 17, 918–926
25. Wu, H., Hu, Z., and Liu, X. Q. (1998) *Proc. Natl. Acad. Sci. U. S. A.* 95, 9226–9231
26. Dupluy, P., Bedouelle, H., Fowler, A., Zabin, I., Saurin, W., and Hofnung, M. (1984) *J. Biol. Chem.* 259, 10606–10613
27. Watanabe, T., Ito, Y., Yamada, T., Hashimoto, M., Sekine, S., and Tanaka, H. (1994) *J. Bacteriol.* 176, 4465–4472
28. Chang, A. C. Y., and Cohen, S. N. (1978) *J. Bacteriol.* 134, 1141–1156
29. Scott, C. P., Abel-Santos, E., Wall, M., Wahnon, D. C., and Benkovic, S. J. (1999) *Proc. Natl. Acad. Sci. U. S. A.* 96, 13638–13643

MEB16   Nco
         Ec

MEB16 [diagram of construct]

M V K V I G R R S L G V Q

Fxa   COOH
        NH2
[circle diagram] GTLENFAEY CFNISTGM
         MBP

[lower left diagram]
E⎡ MBP ⎤
  CFNISTGM
  C
E⎡ MBP ⎤H

MEB16   In vivo splicing products

Ming Fxa 45 kd   (~ 5→10 pmol)
1   2   3   4   5   6   7   8   9   10 lower than 44
X   V   K [I] [G] G [R] [S] S [G]
     K   V   I   R   e   L
            e   s   a   R
                       n

11   12 13 14 15 16 17 18 19 20
[G] V   Q   R   F   F   D   G   G   L
[V]   R   I   I   D   I   L   P

Ming Fxa 42 upper   7→6 pmol
1   2   3   4   5   6   7   8   9   10
G   T   L   E   K   F   A   E   Y   X_?-c?

11   12 13 14 15 16 17 18 19 20
F   N   I   S   T   G   M
                  ———   ——
                 Age I MBP

Ming Fxa 42 lower
1   2   3   4   5   6   7   8   9   10
(M) (F) (N) (I) (S) (T) (G) (M) L   V
 C   K   ate age e   se   A   L   V   I
      n   r   r   L   E   K
                                    F   K

11   12   13   14   15   16   17   18   19   20
I   e   N   G   D   D   G   R   Y   R
     a   I   K   G   K   N   K   K   N
     K   A   N   K   A   K
          A   L
          r   L
          K   e

PM,F2 Sequences

Mime-Version: 1.0
X-Sender: benner@ecori.neb.com
Date: Mon, 20 Sep 1999 11:14:32 +0800
To: Ming-Qun Xu <xum@neb.com>
From: Jack Benner <benner@neb.com>
Subject: F2 Sequences

In vivo MEB16

Ming,
Here are your first two sequences:
F2 upper: may be circular or (blocked)

1	2	3	4	5	6	7	8	9	10
X	V	K	I	I	X	X	Y	S?	L
									R

MV KVI G RR SG.
Ec - M contamination

F2 lower:

1	2	3	4	5	6	7	8	9	10					
C	F	N	I	S	T	G	H	M	K	T	E	E	G	K	L	V	I	W	I

Linear MBP

NdeI

Jack

Ec. | MBP -- En.

Jack S. Benner PhD New England Biolabs
E-mail: benner@neb.com ph 978 927-5054 x366 Fax 978 921-1350
WWW   http://www.neb.com/

(NdeI) MBP
Sequence   *M K I G E G K L V*
                        V

                 I W

Printed for xum@neb.com (Ming-Qun Xu)   1

PAGE 17/18 * RCVD AT 2/24/2006 4:05:18 PM [Eastern Standard Time] * SVR:USPTO-EFXRF-1/2 * DNIS:2738300 * CSID:19783807475 * DURATION (mm-ss):10-32

Project No. _____
Book No. _____   181

Page No. ___

gel / Short / Long / Sequencing   Current: **DATE:** _____

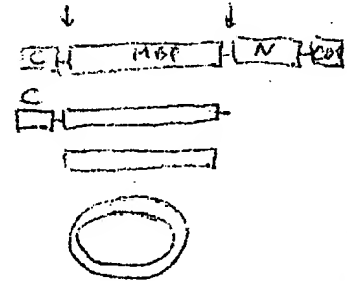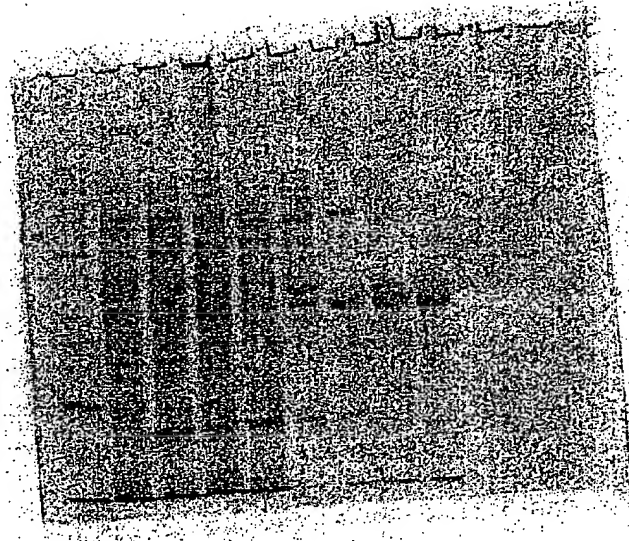__ % gel Agarose / Acrylamide   Power:

is and Conclusions: Voltage:

Reload MEB 16   Time of Run:

Amylose fraction

\+   10mM Tris 100 mM NaCl

1:50 Fxa cleavage 1:2 diluted

10mM of pH?

1 w	10
MEB 16 Uni	2 µl
Ind 15°C	2 µl
Sup	2 µl
Amylose F.T.	2 µl
Amylose F2	1 µl
4°C Fxa 1:100	1 µl
RT Fxa 1:100	1 µl
RT Fxa 1:50	2 µl

To Page No. ____

ssed & Understood by me,	Date	Invented by	Date
		Recorded by	